Nucleic Acids Research, 1997, Vol. 25, No. 10 **1875–1882**

Strong, specific, monodentate G–C base pair recognition by N^7-inosine derivatives in the pyrimidine•purine–pyrimidine triple-helical binding motif

Judith Marfurt, Serge P. Parel and Christian J. Leumann*

Department of Chemistry and Biochemistry, University of Bern, Freiestrasse 3, CH-3012 Bern, Switzerland

Received February 26, 1997; Revised and Accepted April 4, 1997

ABSTRACT

The nucleoside analogs 7-(2′-deoxy-α-D-ribofuranosyl)hypoxanthine (α^7H, 1), 7-(2′-deoxy-β-D-ribofuranosyl)hypoxanthine (β^7H, 2) and 7-(2′-*O*-methyl-β-D-ribofuranosyl)hypoxanthine (β^7H_{OMe}, 3) were prepared and incorporated into triplex forming oligodeoxynucleotides, designed to bind to DNA in the parallel (pyrimidine•purine–pyrimidine) motif. By DNase I footprinting techniques and UV-melting curve analysis it was found that, at pH 7.0, the 15mer oligonucleotides d(TTTTT^{Me}CTXT^{Me}CT^{Me}CT^{Me}CT) (^{Me}C = 5-methyldeoxycytidine, X = β^7H, β^7H_{OMe}) bind to a DNA target duplex forming a H•G–C base triple with equal to slightly increased (10-fold) stability compared to a control oligodeoxynucleotide in which the hypoxanthine residue is replaced by ^{Me}C. Remarkably, triple-helix formation is specific to G–C base pairs and up to 40 μM third strand concentration, no stable triplex exhibiting H•A–T, H•T–A or H•C–G base arrangements could be found (target duplex concentration ~0.1 nM). Multiply substituted sequences containing β^7H residues either in an isolated [d(TTTTTβ^7HTβ^7HTβ^7HTβ^7HT-β^7HT)] or in a contiguous [d(TTTβ^7Hβ^7Hβ^7Hβ^7HTTTT-β^7HTTT)] manner still form triplexes with their targets of comparable stability as the control (^{Me}C-containing) sequences at pH 7.0 and high salt or spermine containing buffers. General considerations lead to a structural model in which the recognition of the G–C base pair by hypoxanthine takes place via only one H-bond of the N–H of hypoxanthine to N7 of guanine. This model is supported by a molecular dynamics simulation. A general comparison of the triplex forming properties of oligonucleotides containing β^7H with those containing ^{Me}C or N^7-2′-deoxyguanosine (N^7G) reveals that monodentate recognition in the former case can energetically compete with bidentate recognition in the latter two cases.

INTRODUCTION

Bidentate thymine-adenine (T•A) and protonated cytosine-guanine (C^+•G) base recognition in the parallel Hoogsteen DNA triple helical motif (py•pu–py motif), (1,2) as well as A•A, T•A and G•G base recognition in the antiparallel reversed Hoogsteen motif (pu•pu–py motif) (3,4) are distinctly favored over other possible base–base combinations (5,6) and form the structural basis of the attractive interaction between a third DNA strand and a DNA duplex. Because of the restriction of both binding motifs to homopurine and homopyrimidine sequences, much effort has recently been devoted to the search for a more general mode of DNA duplex recognition by oligonucleotides. Approaches involved the use of oligonucleotides that were designed to bind to purine–pyrimidine block sequences either exerting the structural properties of both triplex binding motifs in one strand (7–9); or by the use of oligonucleotides that were joined in a 3′–3′ or 5′–5′ direction, and that recognize purine tracts on both strands of the corresponding duplex via the Hoogsteen binding mode (10–14). Other attempts involved the use of oligonucleotides containing modified bases either: (i) to recognize a pyrimidine unit within a purine tract or to span the major groove in order to complex the whole base pair (15–17); or (ii) to bring about less specific interactions at the site of purine–pyrimidine inversion in the duplex (18,19).

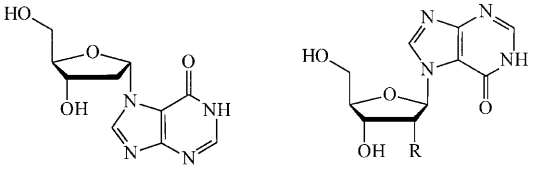
In an effort to search for a new triple-helical binding motif we investigated the triplex formation properties of oligonucleotides containing the non natural nucleosides 7-(2′-deoxy-α- and β-D-ribofuranosyl)hypoxanthine (α^7H) **1** and (β^7H) **2** as well as the 2′-*O*-methyl derivative of the corresponding β-ribonucleoside (β^7H_{OMe}) **3** (Fig. 1). By screening pyrimidine-rich oligonucleotides containing single α- or β^7H residues we found that β^7H (as well as β^7H_{OMe}) and to a lesser extent also α^7H (20) can recognize a G–C base pair with high selectivity and, in the case of β^7H, equal or slightly enhance efficiency compared to 5-methylcytidine at pH 7.0. The stabilizing interaction most likely relies on base-pairing of hypoxanthine to guanine via one H-bond.

MATERIALS AND METHODS

Preparation of oligonucleotides

Oligonucleotides **18**–**25** were synthesized on a 1.3 μmol scale on a Pharmacia Gene-Assembler Special DNA synthesizer using standard β-cyanoethyl phosphoramidite chemistry. Reagents were prepared as described in the user manual and standard

* To whom correspondence should be addressed. Tel: +41 31 631 43 55; Fax: +41 31 631 34 22; Email: leumann@ioc.unibe.ch



1, (α ^{7}H) 2, (β ^{7}H) R = H 3, (β $^7H_{OMe}$) R = OMe

Figure 1. The three N^7-hypoxanthine nucleosides prepared and used in this study.

concentrations of phosphoramidite solutions were used. For the introduction of the unnatural building blocks **7**, **16** and **17**, the standard synthesis cycle was slightly changed to allow for a prolonged (6 min) coupling time. Coupling efficiencies were estimated from trityl assays and were >90% per step for the modified building blocks. After chain assembly (trityl off mode) the oligomers were removed from the solid support and deprotected by standard ammonolysis (25% aq. NH_3, 55°C 16 h). The crude oligomers were purified by DEAE–HPLC and the quality of the isolated material checked by reversed phase HPLC. Purified oligonucleotides were desalted over SEP-PAK C-18 cartridges (Waters) and their structural integrity verified by matrix assisted laser desorption ionization time of flight mass spectrometry (MALDI-TOF-MS). Details of HPLC separation, synthesis and analytical data are with the supplementary material.

T_m measurements

UV-melting curves (λ = 260 nm) were measured on a Varian Cary-3E UV/VIS photospectrometer using a consecutive heating-cooling-heating cycle (0°C–90°C–0°C–90°C) with a linear gradient of 0.5°C/min in the buffer system specified. T_m values were defined as the maximum of the first derivative of the melting curves.

Preparation of plasmid pJM4C1

All DNA and cell manipulations followed standard protocols (21). All aqueous solutions for DNA manipulations were prepared with Milli-Q water (Millipore). *Escherichia coli* XL-1 Blue competent cells were from Stratagene. dNTPs, sonicated calf thymus DNA were from Pharmacia, diluted to the appropriate concentrations and stored at –20°C. [α-^{32}P]dCTP (3000 Ci/mmol, 250 µCi) was from Amersham. Restriction endonucleases, DNase I and DNA polymerase I (Klenow fragment) were from Pharmacia; Ampli*Taq* from Perkin-Elmer; alkaline phosphatase and snake venom phosphodiesterase (crotalus durissus) from Boehringer Mannheim. Buffers and salts (if not provided with the enzymes) were made from Fluka (Microselect grade) or Sigma chemicals. Synthetic oligonucleotides and primers were prepared as described in the 0.2 µmol scale using standard β-cyanoethyl phosphoramidites and solid support (Pharmacia). Modified coupling times of 2 min were used. Oligonucleotides were purified by polyacrylamide gel electrophoresis.

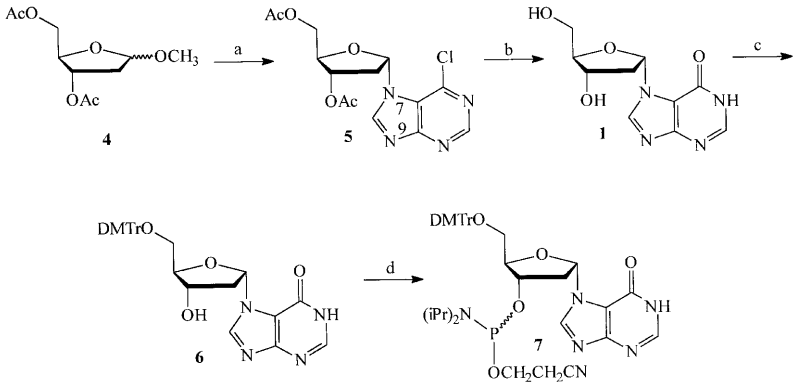
A 164 bp DNA duplex containing the four triplex target cassettes was prepared according to the method of Chen*et al.* (22) from three synthetic oligonucleotides **A**–**C**, partially overlapping by 20 nucleotides, and two primers **P1** and **P2** of the sequences indicated below by polymerase chain reaction (PCR). Oligo **A**: CGACGACCCGGGACGCCGTCGACGCAAAAAGAGAGA-GAGATGCGGGAATTCCGAAAAAGACAGAGAGA, Oligo **B**: TCTTTTTCGGCATGCCCGCATCTCTCTATCTTTTTGCGGA-TCCGGCGTTCTCTCTGTCTTTTTCGGAA, Oligo **C**: TGCGG-GCATGCCGAAAAAGAAAGAGAGAGGCACTGGCCATGA-CCAAGCTTGCGCTCTGCTGAAGCCAG, **P1**: CGACGACCC-GGGACGCCGTC, **P2**: CTGGCTTCAGCAGAGCGCAA. PCR conditions: 10 µl PCR-buffer 10× [100 mM Tris–HCl, pH 8.3, 500 mM KCl, 1.5 mM $MgCl_2$, 0.01% (w/v) gelatine], dATP, TTP, dGTP, dCTP (2 µl of a 10 mM solution of each), 0.1 nmol of primer **P1** and **P2**, 0.1 pmol of each oligonucleotide **A**–**C** and 0.5 µl Ampli*Taq* DNA polymerase (2.5 mU) adjusted to 100 µl volume with H_2O. After adding a layer of nujol, 35 PCR cycles (denaturation: 94°C, 1 min, annealing: 65°C, 1 min, reaction: 72°C, 2 min) were performed and the resulting product purified by 2.5% agarose gel-electrophoresis and isolated from the gel by QIAEXII extraction kit (Qiagen). The 164 nt long duplex was digested with *Hin*dIII and *Ava*I and subsequently ligated into the large *Hin*dIII/*Ava*I restriction fragment of pUC18. *Escherichia coli* XL-1 Blue competent cells were transformed by the ligated plasmid pJM4C1, and plasmid DNA from ampicillin-resistant white colonies was isolated. The presence of the desired insert of the selected clone was determined by forward and reversed Sanger dideoxy sequencing. One mutation (GC→AT) at position 45 of the original 164 nt duplex was detected. The mutation, however, was outside the target cassettes. The plasmid was isolated on a preparative scale using a Maxi Prep kit (Qiagen) to give 640 µg of pJM4C1.

3′ end-labeling of DNA fragment

Approximately 6.4 µg plasmid pJM4C1 were digested with restriction endonucleases *Ava*I and *Pvu*II and the products separated by agarose gel electrophoresis. The desired band was excised and extracted with a QiaexII extraction kit. The restriction fragment was then ethanol precipitated and redissolved in H_2O. 3′ end-labeling was performed with [α-^{32}P]dCTP using the Klenow fragment of DNA polymerase I. Excess unlabeled nucleotide triphosphate (dCTP, dGTP) was added after labeling to ensure complete fill in. Non-incorporated nucleotides were removed by gel filtration over NICK columns (Pharmacia). Finally the labeled fragment was ethanol-precipitated. From 6.4 µg of plasmid a total Cerenkov radioactivity of 2 900 000–3 750 000 c.p.m. was obtained.

DNase I footprint titration

These experiments were performed in analogy to a published protocol (23). Triplex forming oligonucleotides were placed in 18 microcentrifuge tubes to span a concentration range of 0–40 µM (final conc.) and adjusted to a volume of 18 µl with H_2O. To each tube 27 µl of a solution containing the labeled DNA fragment (20 000 c.p.m.) together with 9 µl association buffer 5× [50 mM bis-Tris-HCl, pH 7.0, 500 mM NaCl, 1.25 mM spermine-tetrahydrochloride, sonicated calf thymus DNA (50 µM in base pairs)] were added [final solution conditions: 10 mM bis-Tris–HCl, pH 7.0, 100 mM NaCl, 0.25 mM spermine-tetrahydrochloride, sonicated calf thymus DNA (10 µM in base pairs)]. After 5 days of equilibration at 18°C, 5 µl of a solution containing DNase I (0.54 U) 50 mM $CaCl_2{\cdot}2H_2O$, 50 mM $MgCl_2{\cdot}6H_2O$, 10 µM non- specific DNA single strand d(AATTTAATAT), 10 mM bis-Tris–HCl, pH 7.0



Scheme 1. Reagents and conditions. (a) 6-chloropurine, HMDS, TMSCl, $SnCl_4$, MeCN, r.t., 3 h, 51% (+ N^9-mixture α,β 2.5:1, 37%); (b) 1 M NaOH in THF:MeOH:H_2O 5:4:1, 0–65°C, 5h, 75%; (c) $(MeO)_2TrCl$, pyridine r.t., 3 h, 68%; (d) $(NCCH_2CH_2O)[(iPr)_2N]PCl$, $EtN(iPr)_2$, THF, r.t., 1.5 h, 69%.

and 5% glycerol were added to each probe. After 2 min of incubation at 22°C DNase I activity was quenched by adding 8.3 µl of a solution containing 0.14 M EDTA, 1.4 M NaCl and 0.35 µg/µl glycogen, pH 8.0, followed by 120 µl ethanol. Precipitation was allowed to proceed for 3 h at –20°C and was followed by centrifugation at 16 000 g for 30 min at 4°C. The pellets were washed with 80 µl of 70% ethanol, resuspended in 20 µl H_2O, lyophylized and dissolved in 7 µl of 80% formamide loading buffer. After denaturation (10 min, 85°C) the samples were loaded on an 8% denaturing polyacrylamide gel and separated by electrophoresis in 1× TBE buffer at 1800 V for 75 min. After drying on a slab drier (80°C, 1 h) the gel was exposed to a storage phosphor imager screen (Molecular Dynamics) overnight. The footprinting experiments were all performed in triplicate.

Molecular modeling

Molecular mechanics and dynamics (MD) calculations were performed using the AMBER force field (24) as incorporated in the InsightII (95.0)/Discover (3.0) software package from Molecular Simulations, San Diego, CA, USA. The base triples β-7H•G–C and $^{Me}C^+$•G–C were constructed by modifying the standard T•A–T triple. Partial charges for β^7H were attributed by a donor-acceptor scheme, those for $^{Me}C^+$ were semi-empirically calculated by electrostatic potential fitting (25) using the AM1 model as implemented in the program Spartan (4.0) of Wavefunction, Inc., Irvine, CA, USA. With these building blocks the triplex d($TT^{Me}C^+T\beta^7HT^{Me}C^+T^{Me}C^+$)•d(AAGAGAGAG)–d(CTCT-CTCTT) was constructed using the parameters of a A-conformation triplex (26). Explicit water and counterions were neglected. A distance-dependent dielectric constant of ε = 4r was used to simulate an aqueous environment. No cutoff distance was applied for non-bonded interactions. Van der Waal's and electrostatic 1–4 interactions were scaled by 0.5. No artificial constraints were used. MD were performed in the NVT ensemble keeping the temperature constant by coupling to an external heat bath (27). A time step of 1 fs was used for the numerical integration of the equation of motion.

The triplex was first minimized to a final gradient of 0.05 kcal·mol^{-1}·Å^{-1}. Then the system was equilibrated by progressive heating to 300 K: 1 ps at 50 K and 100 K, 2 ps at 150 K, 3 ps at 200 K, 5 ps at 250 K and 10 ps at 300 K. Finally, the production trajectory of the system was recorded for 200 ps at 300 K, saving instantaneous structures every 0.5 ps. Energy components remained constant during the simulation. Mean values: E_{total} = 1352 kcal·mol^{-1}, $E_{kinetic}$ = 769 kcal·mol^{-1}, $E_{potential}$ = 583 kcal·mol^{-1}.

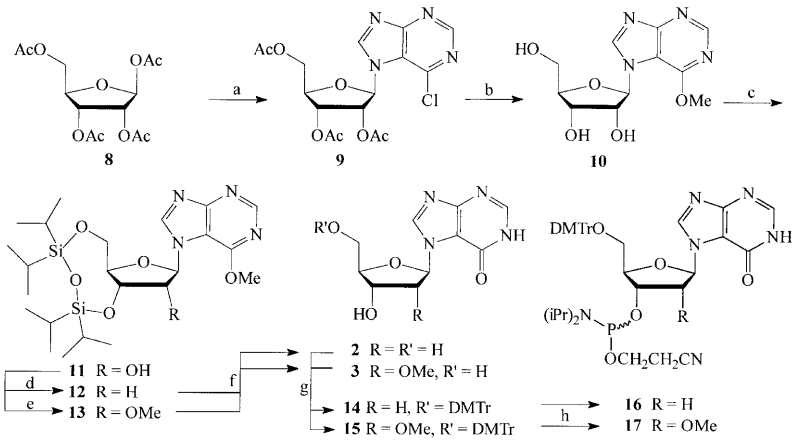
RESULTS

Synthesis of nucleosides and oligonucleotides

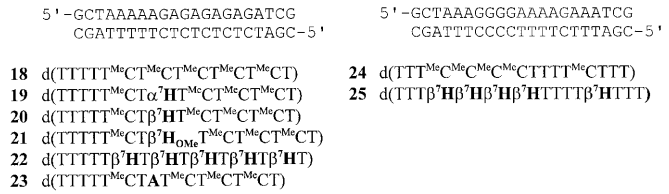
Although N^7α- and β-ribosyl derivatives of hypoxanthine were synthesized before (28–30), oligonucleotides incorporating such units were not reported so far. For this purpose we developed an efficient access to the phosphoramidite building blocks **7**, **16** and **17** of the nucleoside analogs **1**–**3**. Our synthesis of the α-nucleoside **1** started with the readily available 2-deoxyribose derivative **4** and 6-chloropurine and essentially followed the Vorbrüggen procedure for Lewis acid promoted nucleoside synthesis (31). Under optimized conditions (Scheme 1) the α-anomer **5** could be isolated from the reaction mixture as the sole N^7 nucleoside (51%) together with a readily separable 1:1 mixture of the corresponding α,β-N^9-nucleosides (37%). Elaboration of the phosphoramidite **7** was then achieved by standard transformations. The constitution at the anomeric center (α) as well as the point of attachment of the base (N^7) could be verified in the case of the free nucleoside **1** by X-ray structure determination (32).

In order to produce the corresponding nucleosides in the β-series (**2** and **3**), tetraacetyl ribose **8** was used as the starting sugar in the nucleosidation reaction (Scheme 2). Under similar conditions the anomerically pure N^7-β-nucleoside precursor **9** was obtained in 76% yield. The syntheses of **2** and **3** as well as the phosphoramidite building blocks **16** and **17** were then accomplished via a string of standard synthetic transformations, as outlined in Scheme 2. The N^7 base attachement in **2** and **3** was rigorously acertained by comparison of the ^{13}C-NMR chemical shifts of the base carbons with tabulated data (33).

The phosphoramidites **7**, **16** and **17** were subsequently used in the solid-phase oligodeoxyribonucleotide synthesis on a DNA synthesizer. For the study of the triplex forming properties of the modified oligonucleotides, the sequences **18**–**25** containing 1–5 N^7-hypoxanthine residues, together with the two corresponding 21mer DNA duplex target sequences (Fig. 2) were prepared.



Scheme 2. Reagents and conditions. (a) 6-chloropurine, N,O-bis-trimethylsilylacetamide, $SnCl_4$, CH_3CN, 0°C, r.t., 1.5 h, 76%; (b) 0.1 M NaOH in THF:MeOH:H_2O 5:4:1, 0°C, 1 h, 75%; (c) 1,3-dichloro-1,1,3,3-tetraisopropyldisiloxane, pyridine, r.t., 1.5 h, 79%; (d) (i) p-toluoyl chlorothioformate, DMAP, CH_3CN, r.t., 16 h, (ii) Bu_3SnH, AIBN, toluene, 75°C, 2 h, 82%; (e) NaH, MeI, DMF, –25°C, 2.5 h, 82%; (f) 1 M NaOH in THF:MeOH:H_2O 5:4:1, 70°C, 5.5 h, 81% (**2**), 89% (**3**); (g) $(MeO)_2TrCl$, pyridine r.t., 2 h, 54% (**14**), 73% (**15**); (h) $(NCCH_2CH_2O)[(iPr)_2N]PCl$, $EtN(iPr)_2$, THF, 1 h, 63% (**17**), 64% (**16**).



18 d(TTTTT^{Me}CT^{Me}CT^{Me}CT^{Me}CT^{Me}CT)
19 d(TTTTT^{Me}CTα^{7}**H**T^{Me}CT^{Me}CT^{Me}CT)
20 d(TTTTT^{Me}CTβ^{7}**H**T^{Me}CT^{Me}CT^{Me}CT)
21 d(TTTTT^{Me}CTβ^{7}**H**$_{OMe}$T^{Me}CT^{Me}CT^{Me}CT)
22 d(TTTTTβ^{7}**H**Tβ^{7}**H**Tβ^{7}**H**Tβ^{7}**H**T)
23 d(TTTTT^{Me}C**TA**T^{Me}CT^{Me}CT^{Me}CT)
24 d(TTT^{Me}C^{Me}C^{Me}C^{Me}CTTTT^{Me}CTTT)
25 d(TTTβ^{7}**H**β^{7}**H**β^{7}**H**β^{7}**H**TTTTβ^{7}**H**TTT)

Figure 2. The DNA–duplex target sequences and the corresponding triplex forming oligonucleotides prepared and used in this study.

T_m analysis of triplex formation of a α^7H containing oligonucleotide

Recently, Neidle, Reese and coworkers reported on the pairing properties of oligonucleotides containing an α-configured, the base aminopyridine containing nucleoside analog as a replacement for natural cytidine and found it to form triplexes of slightly enhanced stability compared to an oligonucleotide containing the corresponding β-configured analog (34). On the basis of DNase I footprinting experiments we reported that the 15mer oligonucleotide **19** binds with 5–10 fold reduced affinity relative to **18** (containing a ^{Me}C residue) to a triplex target site on a 229 bp DNA duplex under selective formation of a α^7H•G-C base triple (20).

We desired to confirm this result by an independent analytical technique and performed a UV-melting curve analysis of dissociation of oligonucleotide **19** from the corresponding 21mer target DNA duplex containing a G–C base pair opposite the α^7H residue at neutral pH and varying buffer conditions. From the T_m data obtained in the buffer system approaching the conditions of the DNase experiment best, **19** shows reduced affinity to its target by ΔT_m of –10.8°C (Table 1) relative to the control sequence **18**. This compares to a ΔT_m of –19.4°C with the sequence **23** relative to **18**, having an A•G–C mismatch in the center of the triplex. Thus the same trend as in the DNase footprinting experiments was observed; evidence that the decrease in binding efficacy cannot be compared with that of a truly mismatched system. Although triplex binding of third strands with α^7H is inferior in stability compared to that with third strands containing α-aminopyridine nucleosides (34), it is a further demonstration that α-configured nucleosides as components in otherwise β-configured third strands can be used for DNA duplex recognition.

Table 1. T_m data (°C) from UV-melting curves (λ = 260 nm) of third strand dissociation from the corresponding target duplex sequence listed in Figure 2

Oligonucleotide	200 mM NaCl	1 M NaCl	100 mM NaCl, 0.25 mM spermine
18	27.1	30.2	37.7
19	11.7	17.5	26.9
20	27.3	33.2	40.0
21	23.3	28.7	37.5
22	13.1	35.6	38.1
23	~7.0	10.2	18.3
24	11.6	18.2	21.1
25	≤5	22.1	23.6

Concentration of triplex = 1.6 µM, in 10 mM Na-cacodylate, pH 7.0.

DNase I footprinting experiments of β^7H containing oligonucleotides: binding affinity and selectivity

The selectivity of binding of the β^7H and β^7H_{OMe} containing oligonucleotides **20** and **21** was investigated with a DNase I footprint assay according to known methods. Plasmid pJM4C1, containing a DNA insert with the four triplex target cassettes outlined in Figure 3 was prepared and used as a source of target DNA. The 229 bp *Ava*I/*Pvu*II restriction fragment of this plasmid was 3′-end labeled (pyrimidine-rich strand) and incubated with increasing concentrations of oligonucleotide **20** (200 pM–40 µM), in 10 mM bis-Tris–HCl, pH 7.0, 100 mM NaCl, 0.25 mM spermine) at 18°C for 5 days, and then digested with DNase I. The generated reaction products were separated by denaturing gel

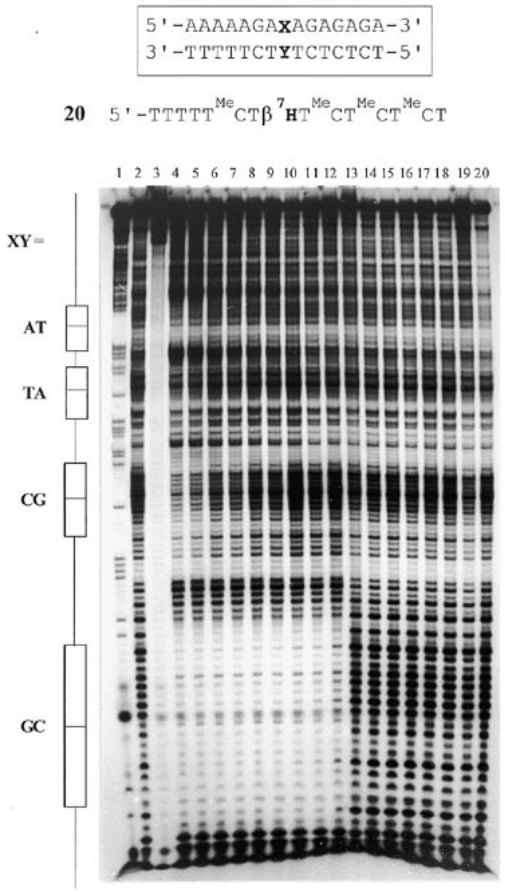


Figure 3. Autoradiogram of an 8% denaturing polyacrylamide gel resulting from a DNase I footprinting experiment of oligonucleotide **20** with the 229 bp DNA fragment that was ^{32}P-labelled at the 3′-end of the pyrimidine rich strand. The boxes to the left indicate the nature of the central base pair and position of the four triplex target cassettes, the sequence of which is indicated on top of the figure. Lane 1, products of an adenine and guanine specific Maxam–Gilbert sequencing reaction; lane 2, DNase I digested duplex; lane 3, intact duplex; lanes 4–20, DNase I digested duplex fragments obtained after incubation with different concentrations of oligonucleotide 20: 40 µM, lane 4; 20 µM, lane 5; 8 µM, lane 6; 4 µM, lane 7; 2 µM, lane 8; 800 nM, lane 9; 400 nM, lane 10; 200 nM, lane 11; 80 nM, lane 12; 40 nM, lane 13; 20 nM, lane 14; 8 nM, lane 15; 4 nM, lane 16; 2 nM, lane 17; 800 pM, lane 18; 400 pM, lane 19; 200 pM, lane 20.

electrophoresis and visualized by storage phosphor autoradiography. Figure 3 shows the autoradiogram of the duplex incubated with oligonucleotide **20**. From this it becomes clear that **20** binds to the G–C cassette already at 80 nM third strand concentration. Triplex formation is highly specific. No binding to any of the other cassettes is observed within the screened concentration range. The same is also true for oligonucleotide **21**, in which β^7H is replaced by β^7H_{OMe} (data not shown). We have calculated equilibrium association constants (K_a) for binding of oligonucleotides **18**, **20** and **21** to the G–C containing cassette as described (23) and found them to amount to 7.4 (± 3.0) × 10^6 M^{-1} for oligo **18**, 6.7 (± 2.0) × 10^7 M^{-1} for oligo **20** and 2.6 (± 2.4) × 10^6 M^{-1} for oligo **21**. This indicates an ~10-fold (average) increase of binding of the β^7H containing oligo **20** relative to the ^{Me}C containing control **18** and a 5-fold decrease in affinity of the β^7H_{OMe} containing oligo **21** relative to **18** in the buffer described.

T_m analysis of β^7H containing oligonucleotides: multiple substitutions and sequence dependence

Melting analysis of the triplex at neutral pH with oligonucleotide **20**, containing one β^7H residue opposite a G–C base pair gives T_m values that are slightly enhanced to that obtained from the triplex containing the reference oligonucleotide **18** with ^{Me}C opposite to G–C of the duplex at pH 7.0 (Table 1). Oligo **21**, having a β^7H_{OMe} unit in the center of the sequence, displays equal to slightly reduced affinity to the duplex target, depending upon the buffer used. Thus T_m analysis and DNase I footprint assay are in good agreement with respect to the relative stabilities of the three oligonucleotides.

Again at pH 7.0, sequence **22** in which all five ^{Me}C-residues were replaced by β^7H units shows slightly increased stability relative to the control **18** at high salt (1 M NaCl) or spermine containing buffers, while its affinity to the target is decreased at low salt (200 mM NaCl). As expected pairing of **22** to its target duplex is pH independent in the range of 6.0–8.0 (T_m = 35.0 ± 1.1°C, 1 M NaCl, standard triplex concentration). The relative destabilization of **22** relative to **18** at low salt is obviously due to the fact that no base in **22** is protonated in the triplex. Thus interstrand charge repulsion between the duplex and the third oligonucleotide, which is less pronounced in the case of **18** compared to **22,** leads to the energetic advantage of the ^{Me}C-containing triplex at low salt concentration.

In order to test the sequence dependence in triplex formation we determined T_m data for the melting of **25**, that contains four contiguous β^7H units from the corresponding 21mer DNA target duplex and compared it to that of **24**, in which all β^7H residues were replaced by the standard ^{Me}C at pH 7.0. At high salt the affinity of the β^7H-containing **25** was found to be superior than that of the control **24**. Again, at low salt the relative affinities are reversed, favoring the ^{Me}C containing sequence most likely for the same reason as in the former sequence system. Melting and subsequent reassociation curves for sequences **18**/**22** and **24**/**25** with their target DNA duplex as an example are depicted in Figure 4. De- and renaturation curves for the third strands (transition at lower temperature) in the ^{Me}C containing sequences **18** and **24** are not superimposable while they are in the case of **22** and **25**. This is an indication that triple helix formation kinetics is faster in the β^7H-containing oligonucleotides.

Modeling of a triplex containing a β^7H•G–C base triple

The β^7H containing oligonucleotides bind in a parallel fashion to the purine duplex. This reduces the number of possible base-triples to be taken into consideration. Furthermore, the lactam form in hypoxanthine as well as in N^7 or N^9 derivatives thereof is by far the most stable tautomeric form (33). Although a syn conformation of the base relative to the sugar is formally possible, the base must adopt an anti orientation in order to form any H-bond with the acceptor purine site. By applying these criteria there remain two possible base arrangements for the triplex that are displayed in Figure 5 (top). They diverge from each other by the nature of the acceptor site of the H-bond, either being N^7 or O^6 of guanine.

We have modeled the nonamer triple helix d(TT^{Me}C$^+$Tβ^7HT-^{Me}C$^+$T ^{Me}C$^+$)•d(AAGAGAGAG)–d(CTCTCTCTT) containing a central β^7H•G–C base-triple. This sequence corresponds exactly to the internal part of oligonucleotide **20** and its target. Due to the fact that there exists no unique conformational model

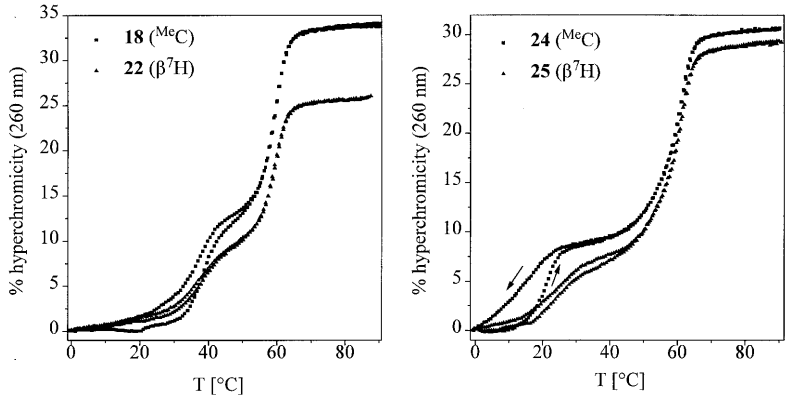


Figure 4. UV-melting curves (heating and cooling cycle) of oligonucleotides **18**, **22**, **24** and **25** (concentration = 1.2–1.6 µM in 10 mM Na-cacodylate, 100 mM NaCl, 0.25 mM spermine, pH 7.0) with the corresponding DNA target 21mers shown in Figure 2.

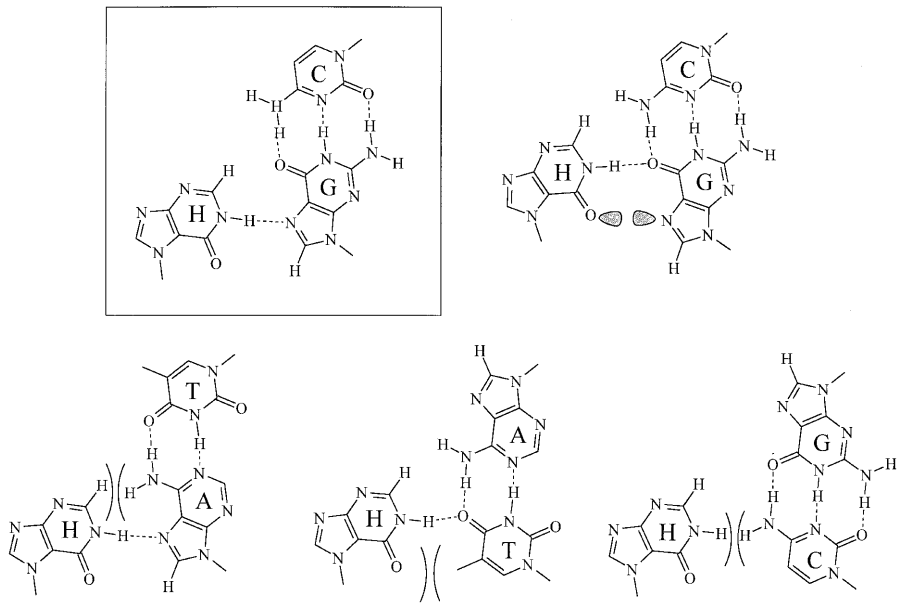


Figure 5. Base arrangements of the N^7-hypoxanthine residues with the natural DNA Watson–Crick base pairs.

for DNA triplexes, we arbitrarily chose and constructed a standard A-DNA triple helix with the N–H of the base hypoxanthine located between N^7 and O^6 of guanine, forming a bifurcated H-bond to both centers in the starting structure. After energy minimization (AMBER force field) an unconstrained molecular dynamics simulation on a 200 ps trajectory was performed and structures were sampled every 0.5 ps. Figure 6 shows a stereoview of the average of the last 100 structures (Fig. 6a) and a detailed view of the base triple formed between β^7H, G and C (Fig. 6b). From the simulation it becomes clear that the triplex is stable throughout the whole 200 ps. Furthermore there exists a firm H-bond between N–H of hypoxanthine to N^7 and not O^6 of guanine. This is best illustrated by the distance versus time plots of NH–N^7 and NH–O^6 during the simulation (Fig. 7). From these plots an average distance of 2.21 Å for NH–N^7 and 2.66 Å for NH–O^6 was calculated favoring clearly the N^7 centered H-bond over the bifurcated or O^6-centered H-bond. Empirically one can explain this preference by the expected electrostatic repulsion between the O^6 of hypoxanthine and N^7 of guanine, and eventually by Van der Waal's repulsions between H–C2 of hypoxanthine and H–N^4 of cytosine in the alternative model (Fig. 5, top).

DISCUSSION

Base-pairing versus intercalation

There exists the formal possibility that the base hypoxanthine as part of a third strand could intercalate in the target duplex rather than forming a base-triple. Precedence for this is found in the case of the artificial base D3 that was designed to bind to a C–G base pair (15), but was shown to intercalate in the duplex (35). The fact that sequences with multiple substitution of β^7H as **22**, and even **25**, containing consecutive β^7H residues, still interact with the corresponding target duplexes in comparable stability as the ^{Me}C

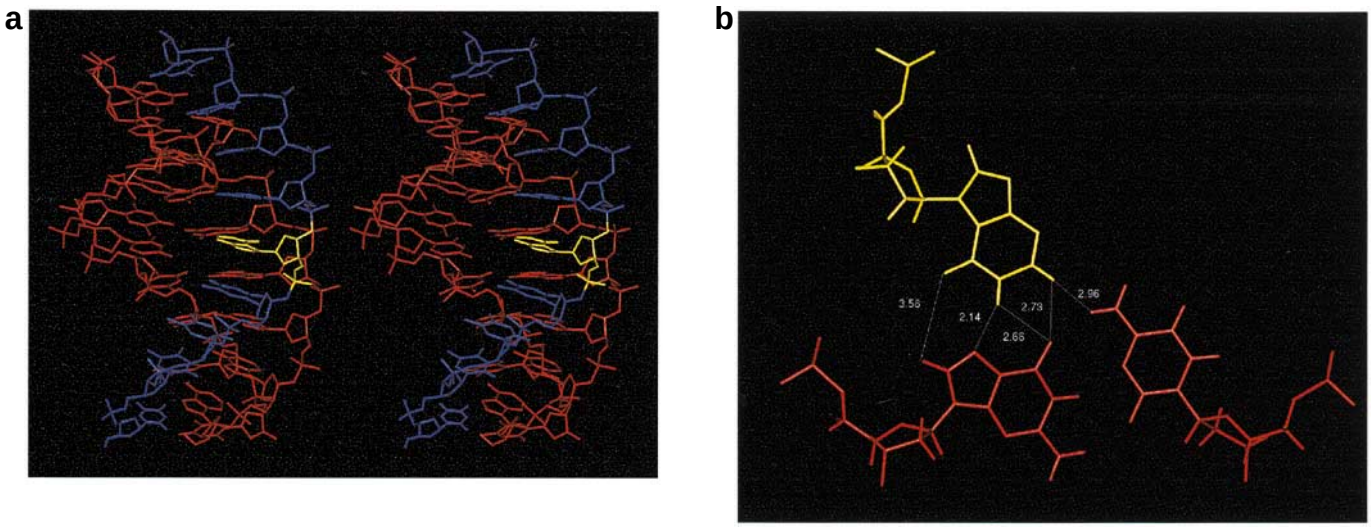


Figure 6. (**a**) Stereoview of the average structure of the last 50 ps from the 200 ps dynamics simulation of the triplex d(TT^{Me}C^{+}Tβ^{7}HT^{Me}C^{+}T ^{Me}C^{+})•d(AAGAGAGAG)–d(CTCTCTCTT). The Watson–Crick duplex is colored in red, and the third strand in blue with the β^{7}H residue in yellow. (**b**) Detailed view of the corresponding β^{7}H•G–C base-triple with selected distances (Å).

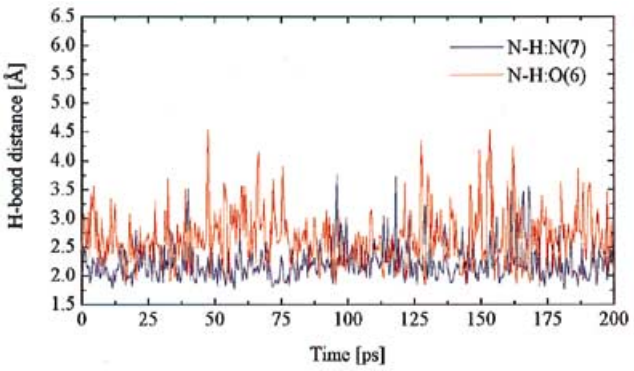


Figure 7. Distance versus time plot of N1–H of hypoxanthine to N7 (red) and O6 (blue) of guanine within the base-triple β^{7}H•G–C.

containing control sequences, together with the fact that base pair recognition is specific, almost certainly rules out any other mode of interaction than base–base recognition via a specific H-bonding scheme.

Selectivity and stability of base-triple formation

The observed striking specificity for the β^{7}H•G–C base-triple raises the question about its origin. If one assumes the N1–H···N7 model (Fig. 5 top, left) to be the structurally and energetically relevant one within the given triplex, then the other possible base arrangements can be constructed by mutation of the isomorphic Watson–Crick base pairs of the target duplex (Fig. 5 bottom). This simple picture offers an answer to the G-selectivity in the recognition process in that all alternatives suffer at least from one repulsive Van der Waal's interaction.

The N^{1}–H···N^{7} H-bond therefore seems to be a necessary requirement in H•G–C recognition. However it should be noted that additional energetic contribution may arise from enhanced intrastrand stacking around the modified bases. Likewise, it cannot be excluded that favorable electrostatic interactions between H–C2 of hypoxanthine and O^{6} of guanine help stabilizing the β^{7}H•G–C base-triple.

Comparison with N^{7}G in triple helix formation

Hunziker and Dervan (36), and Brunar and Dervan (37) investigated the non-natural nucleoside N^{7}-2′-deoxyguanosine (N^{7}G) as a cytidine replacement in triple helix forming oligonucleotides. On the basis of DNase footprint experiments these authors demonstrated convincingly that N^{7}G functions at neutral pH by specifically binding to a G–C base pair with comparable efficiency as ^{Me}C in monomodified sequences. In the context of multiply modified oligonucleotides it was shown that sequences having a large number of GA and AG steps (corresponding to isolated N^{7}G residues in the third strand) are less efficiently bound by N^{7}G whereas sequences with a low number of AG or GA steps (corresponding to blocks of consecutive N^{7}G nucleosides in the third strand) bind to their duplex target with higher affinity than ^{Me}C. Differential stability as a function of the number of GA to AG inversions in the sequence was interpreted as a consequence of the lack of structural isomorphism of a N^{7}G•G–C versus T•A–T base-triple.

β^{7}H can be considered as a derivative of N^{7}G, in which the 2-amino function is lacking. Considering isomorphous base-triple formation in both cases (Fig. 8), the following two observations are of relevance: first, the 2-amino-function in N^{7}G is not essential for specific recognition of the G–C base pair. Second, its presence leads not only to attractive, but also to repulsive effects in the triplex. In sequences containing multiple isolated N^{7}G residues, the presence of the 2-amino group destabilizes the triplex relative to β^{7}H whereas a stabilizing effect is observed for sequences containing multiple consecutive N^{7}G residues. This can be deduced from the relative stabilities of the N^{7}G and β^{7}H sequences compared with those containing ^{Me}C at neutral pH under comparable buffer conditions.

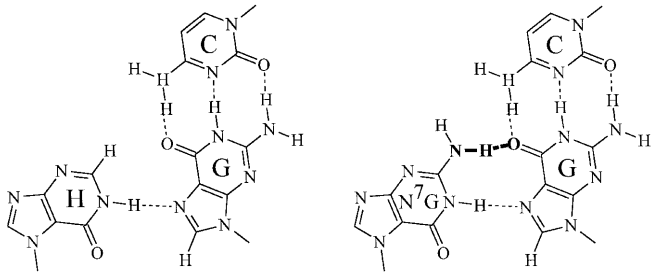


Figure 8. Structures of the β^7H•G–C base-triple and the isomorphous N^7G•G–C base-triple.

Comparison with known monodentate third-strand base contacts

A number of remarkably stable, monodentate third-strand base–base contacts are known, the most prominent examples being the G•T–A triple in the pyrimidine motif (5,38) and the T•C–G triple in the purine motif. (6) However, in all cases the affinity of the corresponding bases to these secondary binding sites is lower than that to their primary target, to which they bind via two H-bonds. Therefore multiple occupation of the secondary binding sites by a sequence-designed oligonucleotide is unlikely to occur in terms of affinity and also selectivity since the corresponding bases in such third strands will preferentially target their primary binding sites.

Oligonucleotides containing non-DNA bases, designed to bind to pyrimidines by one H-bond were also reported. 2′-deoxynebularine, containing a N^9-linked unsubstituted purine heterocycle was shown to bind to C–G pairs in the purine motif (16). But again, C–G recognition is compromised by reduced affinity and reduced selectivity relative to the conventional base triplets. In this light the monodentate recognition of G–C by β^7H is exceptional in both stability and selectivity.

In conclusion, we demonstrated here that up to one third of all base residues in a triplex forming oligonucleotide can be replaced by a heterocycle that recognizes a base in a target duplex by only one H-bond without losing essential binding energy relative to known, stable models that all rely on bidentate base-base recognition schemes. We feel that this finding might refuel investigations directed to the construction of bases that selectively and efficiently recognize pyrimidine bases, thus aiming at the generalization of the triplex approach to DNA duplexes of any base-sequence.

ACKNOWLEDGEMENTS

We thank the Swiss National Science Foundation (Grant 20-42107.94), The Wander-Stiftung, Bern and Ciba-Geigy AG, Basel for generous financial support. We also thank Dr J. Hunziker for very valuable experimental advice and fruitful discussions.

SUPPLEMENTARY MATERIAL

Synthetic details and spectral data for compounds **1**–**17**, together with HPLC- and matrix assisted laser desorption ionization time of flight (MALDI-TOF) MS-analytical data of oligonucleotides **18**–**25**. This material is contained in the internet form of the journal.

See supplementary material available in NAR Online.

REFERENCES

1 Moser,H.E. and Dervan,P.B. (1987) *Science*, **238**, 645–650.
2 François,J., Saison-Behmoaras,T. and Hélène,C. (1988) *Nucleic Acids Res.* **16**, 11431–11440.
3 Durland,R.H., Kessler,D.J., Gunnell,S., Duvic,M., Pettitt,B.M. and Hogan,M.E. (1991) *Biochemistry*, **30**, 9246–9255.
4 Beal,P.A. and Dervan,P.B. (1991) *Science*, **251**, 1360–1363.
5 Best,G.C. and Dervan,P.B. (1995) *J. Am. Chem. Soc.* **117**, 1187–1193.
6 Greenberg,W.A. and Dervan,P.B. (1995) *J. Am. Chem. Soc.* **117**, 5016–5022.
7 Sun,J.S., De Bizemont,T., Duval-Valentin,G., Montenay-Garestier,T. and Hélène,C. (1991) *C. R. Acad. Sci. Paris*, **t. 313, Série III**, 585–590.
8 Beal,P.A. and Dervan,P.B. (1992) *J. Am. Chem. Soc.* **114**, 4976–4982.
9 De Bizemont,T., Duval-Valentin,G., Sun,J., Bisagni,E., Garestier,T. and Hélène,C. (1996) *Nucleic Acids Res.* **24**, 1136–1143.
10 Froehler,B.C., Terhorst,T., Shaw,J. and McCurdy,S.N. (1992) *Biochemistry*, **31**, 1603–1609.
11 Horne,D.A. and Dervan,P.B. (1990) *J. Am. Chem. Soc.* **112**, 2435–2437.
12 Ono,A., Chen,C. and Kan,L. (1991) *Biochemistry*, **30**, 9914–9921.
13 McCurdy,S., Moulds,C. and Froehler,B. (1991) *Nucleosides Nucleotides*, **10**, 287–290.
14 Zhou,B.-W., Marchand,C., Asseline,U., Thuong,N.T., Sun,J., Garestier,T. and Hélène,C. (1995) *Bioconjugate Chem.* **6**, 516–523.
15 Griffin,L., Kiessling,L.L., Beal,P.A., Gillespie,P. and Dervan,P.B. (1992) *J. Am. Chem. Soc.* **114**, 7976–7982.
16 Stilz,H. and Dervan,P.B. (1993) *Biochemistry*, **32**, 2177–2185.
17 Huang,C., Cushman,C.D. and Miller,P.S. (1993) *J. Org. Chem.* **58**, 5048–5049.
18 Kandimalla,E.R., Manning,A.N., Venkataraman,G., Sasisekharan,V. and Agrawal,S. (1995) *Nucleic Acids Res.* **23**, 4510–4517.
19 Durland,R.H., Rao,T.S., Bodepudi,V., Seth,D.M., Jayaraman,K. and Revankar,G.R. (1995) *Nucleic Acids Res.* **23**, 647–653.
20 Marfurt,J., Hunziker,J. and Leumann,C. (1996) *Bioorg. Med. Chem. Lett.* **6**, 3021–3024.
21 Sambrook,J., Fritsch,E.F. and Maniatis,T. (1989) *Molecular Cloning—A Laboratory Manual*. Cold Spring Harbor Laboratory Press,
22 Chen,G.-Q., Choi,I., Ramachandran,B. and Gouaux,J.E. (1994) *J. Am. Chem. Soc.* **116**, 8799–8800.
23 Priestley,E.S. and Dervan,P.B. (1995) *J. Am. Chem. Soc.* **117**, 4761–4765.
24 Weiner,S.J., Kollman,P.A., Nguyen,D.T. and Case,D. (1986) *J. Compt. Chem.* **7**, 230.
25 Cox,S.R. and Williams,D.E. (1981) *J. Compt. Chem.* **2**, 304.
26 Arnott,S. and Selsing,E. (1974) *J. Mol. Biol.* **88**, 509–521.
27 Berendsen,H.J.C., Postma,J.P., van Gunsteren,W.F., DiNola,A. and Haak,J.R. (1984) *J. Phys. Chem.* **81**, 3684.
28 Seela,F. and Winter,H. (1995) *Nucleosides Nucleotides*, **14**, 129–142.
29 Rousseau,R.J., Robins,R.K. and Townsend,L.B. (1968) *J. Am. Chem. Soc.* **90**, 2661–2668.
30 Montgomery,J.A. and Thomas,H.J. (1969) *J. Org. Chem.* **34**, 2646–2650.
31 Vorbrüggen,H. and Bennua,B. (1981) *Chem. Ber.* **114**, 1279–1286.
32 Marfurt,J., Stulz,E., Trafelet,H.U., Zingg,A., Leumann,C., Hazenkamp,M., Judd,R., Schenker,S., Strouse,G., Ward,T.R., Förtsch,M., Hauser,J. and Bürgi,H. (1996) *Acta Crystallogr., Sect. C*, **C52**, 713–716.
33 Chenon,M.-T., Pugmire,R.J., Grant,D.M., Panzica,R.P. and Townsend,L.B. (1975) *J. Am. Chem. Soc.* **97**, 4636–4642.
34 Bates,P.J., Laughton,C.A., Jenkins,T.C., Capaldi,D.C., Roselt,P.D., Reese,C.D. and Neidle,S. (1996) *Nucleic Acids Res.* **24**, 4176–4184.
35 Koshlap,K.M., Gillespie,P., Dervan,P.B. and Feigon,J. (1993) *J. Am. Chem. Soc.* **115**, 7908–7909.
36 Hunziker,J., Priestley,E.S., Brunar,H. and Dervan,P.B. (1995) *J. Am. Chem. Soc.* **117**, 2661–2662.
37 Brunar,H. and Dervan,P.B. (1996) *Nucleic Acids Res.* **24**, 1987–1991.
38 Griffin,L. and Dervan,P.B. (1989) *Science*, **245**, 967–971.